

# 視覚フィードバックが学習者の英語発音に与える影響

野本 尚美・平塚紘一郎

(2017年3月10日受理)

## The Effects of Visual Feedback on Learners' English Pronunciation

Naomi NOMOTO・Kouichirou HIRATSUKA

### 1. 研究背景

実践的コミュニケーション能力の習得を目指す近年の英語教育において、音声指導に対する関心が著しく高まってきている。ベネッセ教育総合研究所が実施した『中高の英語指導に関する実態調査2015』によれば、中学校・高校の英語教員が授業での指導方法として「よく行う」「ときどき行う」と答えた割合が最も高かったものが「音読」、次点が「発音練習」であった。これら二つの指導が「文法の説明」よりも上位であったことは、従来文法偏重主義と言われてきた日本の英語教育が大きな転換期を迎えていることを示唆している。また『小学生の英語学習に関する調査』（ベネッセ教育総合研究所, 2015）においても、外国語活動として7割以上の子どもが「英語のことは（cat, appleなど）を言う練習」「英語の発音練習」「短い文や質問を英語で言う練習」を「している」「いつもしている」＋「ときどきしている」と回答しており、小学校においても英語の音声に慣れ親しむことを目的とした活動が多く取り入れられていることがわかる。

日本人英語学習者による発話音声に関しては、これまでいくつかの実践的研究が行われている。Yabuuchi and Sato(2001)は、日本人英語学習者による英文の音読を10人の英語母語話者に評価させ、そのF0 (fundamental frequency: ピッチ) や、休止の長さなどについて分析した。その結果、F0の最大値及び幅が大きい音読や、文と文

の間の休止が節やフレーズ間の休止よりも長かった音読が高い評価を得たことから、シャドーイングなどを音声指導として取り入れることを提案している。また、発音と理解度 (intelligibility) との関連を調べた研究としてKashiwagi & Snyder (2010) は、日本人英語学習者が発話した音声をアメリカ人と日本人英語教師それぞれに書き取らせ、どのような発音的特徴が理解度に影響を与えるのか調査している。母語話者が聞き手の場合は母音の間違いが理解度を低下させ、また日本人が聞き手の場合は正しい母音の発音に加え発話スピードがやや遅く、ピッチ幅がある方がより聞き取りやすいと結論付けている。これらの研究から、日本人英語学習者が習得すべき音韻やプロソディーについては徐々に明らかになってきたものの、教育現場におけるそれらの具体的な指導法やその効果についてはほとんど検証されていない。

近年の情報通信機器の発達により、発音を視覚的に捉えることが可能となり、それを指導に活用する試みも行われている。Le & Brook (2011) は、母語話者と学習者のピッチ波形を音声分析フリーウェアであるPraat (Boersma & Weenink, 2009) によって提示し、イントネーションを矯正する指導法を提案している。またWilson (2008) は、Praatによる分析画面の見方を学生に教え、学習管理システムのMoodle (Dougiamas & Taylor, 2003) を通してその結果を報告させることにより、発音指導が特に必要な学習者の選定や

学習進捗状況の管理が容易になると述べている。しかしこれらの研究においては詳しい音声分析が行われていないため、視覚フィードバックを用いた発音指導の有効性は実証されていない。またPraatのインターフェースは英語である上、専門用語が多く、日本人英語学習者がPraatを直接操作することは容易ではないことから、これらの研究を日本人英語学習者に対する実践的指導の中で応用するのは難しいと考えられる。

そこで本研究では、PraatとMoodleを連携させた英語発音学習システムを開発し、学習者が容易にPraatの分析結果を参照できる環境を設定した。Praatによって視覚的に捉えることができる情報としては、ピッチやインテンシティー、フォルマントなどが挙げられるが、本研究ではいくつかの先行研究において言及されているピッチに焦点を当て、ピッチ曲線を学習者に提示することの有効性について調査した。

## 2. 研究課題

本研究では英語母語話者によるモデル音声と学習者の発話音声を参照する聴覚フィードバック（Auditory Feedback）グループ（以下AFグループ）と、モデル音声や学習者の発話音声だけでなくピッチ曲線を参照する視聴覚フィードバック（Audio-Visual Feedback）グループ（以下AVFグループ）を比較することにより、視覚フィードバックが学習者の発音に与える影響について分析する。本研究の研究課題は以下の通りである。

- (1) AFグループとAVFグループでは、学習者による発話音声のピッチ幅に違いが見られるのか。
- (2) AFグループとAVFグループでは、学習者による発話音声の1秒当たりのピッチ変化量に違いが見られるのか。

前述したように、先行研究においては発話音声のピッチ幅（max pitchとmin pitchの差）が大きい方が、評価が高い（または理解されやすい）と言われているが、ピッチ幅だけでは「学習者がどれくらい抑揚をつけて発話しているか」という点について計測することはできない。そこで本研究

では、PraatのPitch listing機能（一定時間ごとのピッチの値がリストアップされる機能）を用いて発話音声のピッチ変化総量を求め、それを発話時間で割ることで、1秒当たりのピッチ変化量を算出し、発話音声の抑揚を表す指標として用いた。

## 3. 英語発音学習システムの開発

### 3.1 設計

Praatは様々な音声処理を行うことができ、その解析結果を数值的、視覚的に表示することができる一方、先述したように学習者に直接操作させることは困難である。そこで本研究では、Praatをバックエンドとして動作させて音声解析を行うことにした。Praatはスクリプトによる動作も可能であり、GUIを用いなくても音声解析を行うことが可能である。フロントエンドとしてはLMSの1つであるMoodleを用いた。Moodleはインターフェースを日本語にできるため、日本人学習者にとって操作しやすい。また学習者はファイルを容易に提出することができるため、教員は個別のファイルを参照し、学習管理を行うことができる。以上のような利点から、本研究ではPraatとMoodleを連携させた英語発音学習システムを開発した。

### 3.2 構成

本研究で用いた英語発音学習システムの構成を図1に示す。

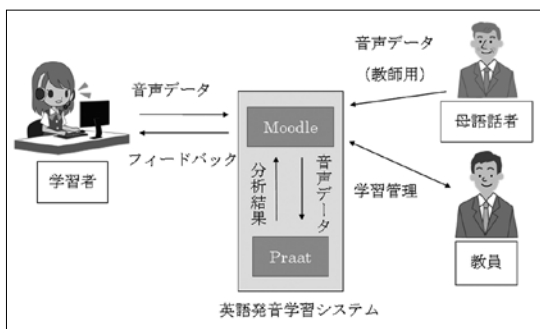


図1. 英語発音学習システムの構成

本システム利用の流れとして、まず教員がMoodle

上にモデルとなる音声を教師用データとして掲載する。次に、学習者はMoodleを使い、ヘッドフォン等で掲載された教師用音声データを参照しながらマイクで自身の音声を録音する。録音された学習者の音声はMoodleからPraatへと送られ、分析が行われる。Praatによる分析結果はMoodleに表示されるようになっており、学習者は視覚フィードバックを参照することができる。学習者はそのフィードバック（ピッチ曲線や数値など）を見て、再度録音を行う。以上を繰り返し、最終的な英語発話音声をMoodleへ提出させる。その後、教員が提出された音声を評価する。

図2は本研究で用いた視覚フィードバックである。モデル音声データとそのピッチ情報（波形と数値）を掲載している。左上のボタンで学習者は録音を行う。



図2. システム画面（録音前）

図3は学習者が音声を録音した後、Praatによる分析結果から視覚フィードバックがMoodleへ表示された画面である。モデル音声データのピッチ波形の下に、学習者が録音した音声のピッチ波形と数値が表示される。学習者はこれらのフィードバックをもとに何度でも録音を行うことができる。

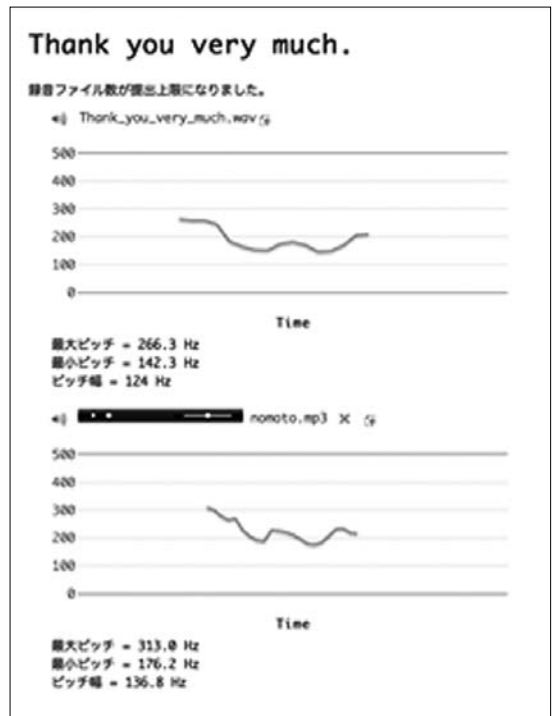


図3. システム画面（録音後）

## 4. 研究方法

### 4.1 研究協力者と調査対象とした英語音声

本研究における研究協力者は非英語専攻の30名の日本人女子短期大学生（19歳～20歳）である。15名ずつの2グループに分け、実用英語技能検定のリーディング・リスニング問題を解答させた結果、グループ間において統計的有意差は認められなかったため、2グループはリーディング、リスニングの習熟度に違いがないと判断した。

中川（2002）による『日本人学生による読み上げ英語音声データベース（以下UME-ERJ）』の中に録音されている英文の中から、日本人学習者にとって特に馴染み深く、実用的なフレーズとして“Thank you very much.”を選び、学習者に発話させた。実験中で用いたモデル音声は、UME-ERJの中のアメリカ人女性の音声をを用いた。

### 4.2 調査手順

学生一人につき一台ずつノートパソコン及びヘッドセットを用意し、各自パソコン上でMoodleの画面を開かせ、以下の手順で発音練習

をさせた。

まず教師用パソコンからスピーカーを通してモデル音声を流し、そのあと学習者全員にリピートさせる一斉発音練習を5回行った。その直後にAudacity（オーディオ編集ソフトウェア）を用いて各自のパソコン及びヘッドセットで自分の発音を録音させ、Moodle上に提出させた。この音声本研究ではpre-testとした。

AFグループにおいては、モデル音声をあらかじめMoodle上に掲載し、学生が各自ヘッドセットを用いて聞けるように設定した。学生に対しては、モデル音声と自分の発話音声を繰り返し再生して発音練習ができることを日本語で説明した。一方、AVFグループにおいては、前章にて述べたように、音声だけでなくピッチ曲線及びピッチに関する情報もMoodle上に掲載し、学生がモデル音声と自分の発話音声について視覚フィードバックも参照しながら発音の練習ができるよう設定した。学生に対しては、モデル音声と自分の発話音声の再生方法に関する説明に加え、ピッチ曲線やピッチに関する数値についても簡単な説明を日本語で行った。どちらのグループにおいても発音に関する指導は一切行わなかった。この個別発音練習を各グループにてそれぞれ10分間行い、その1週間後、もう一度同じ練習を10分間行った上で、自分が最も上手く発音できたと思う音声をMoodle上に提出させた。この音声をpost-testとし、ピッチ幅と1秒あたりのピッチ変化量についてPraatを用いて分析した。

## 5. 結果

### 5.1 ピッチ幅

各グループにおけるピッチ幅は表1の通りである。図4に示すように、両グループ間には差がほとんど見られなかった。AVFグループではpost-testにおいてやや平均値が上昇したものの、二元配置分散分析を行った結果、両グループ内におけるpre-testとpost-test間、またpost-testにおける両グループ間に有意差は認められなかった。

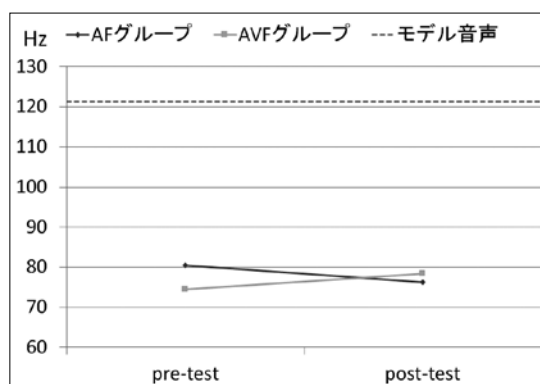


図4. ピッチ幅の推移

### 5.2 1秒あたりのピッチ変化量

各グループにおける1秒あたりのピッチ変化量は表2の通りである。図5の推移を見てみると、AVFグループはpost-testにおいて変化量をやや増加させたものの、二元配置分散分析の結果、両グループ内のpre-testとpost-test間、またpost-testにおける両グループ間に有意差は認められなかった。

しかし、学習者一人一人の値に注目すると、AVFグループにおいては15人中11名（73.3%）の学生がpre-testよりもpost-testにおいて変化量を増加させていることがわかった。一方で、AFグループにおいては変化量が増加した学生は15人中6人（40.0%）にとどまった。

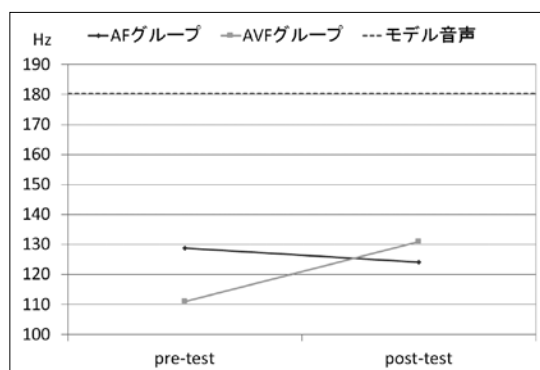


図5. 1秒あたりのピッチ変化量の推移

表1. 各グループにおけるピッチ幅 (単位: Hz)

		平均値	最小値	最大値	標準偏差
AFグループ (n=15)	pre-test	80.5	27.4	162.4	30.5
	post-test	76.3	34.8	110.7	19.1
AVFグループ (n=15)	pre-test	74.5	20.9	118.3	19.8
	post-test	78.4	25.1	175.2	36.5

表2. 各グループにおける1秒あたりのピッチ変化量 (単位: Hz)

		平均値	最小値	最大値	標準偏差
AFグループ (n=15)	pre-test	128.9	75.1	186.6	27.5
	post-test	124.1	58.5	180.4	35.3
AVFグループ (n=15)	pre-test	111	13.9	165.9	33.4
	post-test	131.1	24.2	208.3	40.8

## 6. 考 察

前章の結果を踏まえ、本研究において設定した課題について以下のように述べるができる。

(1) AFグループとAVFグループでは、学習者による発話音声のピッチ幅に違いが見られるのか。

グループ内、グループ間において有意差は認められなかったため、ピッチ幅に違いは見られなかった。

(2) AFグループとAVFグループでは、学習者による発話音声の1秒当たりのピッチ変化量に違いが見られるのか。

1秒当たりのピッチ変化量においても、グループ内、グループ間で有意差は認められなかった。しかしAVFグループにおいては、約7割の学生がpost-testにおいて1秒当たりのピッチ変化量を増加させたことから、視覚フィードバックが学生の発話音声に影響を及ぼした可能性も示唆される。

本研究において視覚フィードバックの効果を明確に実証できなかった理由として、以下の2つが考えられる。まず、練習時間と回数が非常に限られていたという点が挙げられる。本研究では10分間の練習を2回行ったが、ヘッドセットやパソコンの操作に慣れていない学生が発音練習にあまり集中することができずに音声を提出した可能性も考えられる。視覚フィードバックの効果について明らかにするには、さらに長期的に調査を行う必要があると考える。もう一つの問題点として、システムの利便性が挙げられる。今回開発したシス

テムでは、ピッチ曲線を表示するスペースが限られていたため、モデル音声のピッチ曲線と自分の声のピッチ曲線を見比べて練習することが困難であったと考えられる。視覚フィードバックの有効性を調査するためには、システムをさらに改善し、学習者にとって参照しやすい提示方法を確立することも必要である。

## 7. 今後の展望

MoodleとPraatを連携させた英語発音学習システムを用いて、視覚フィードバックが日本人英語学習者の発話に与える影響について調べた結果、本研究では視覚フィードバックの効果は実証されなかった。しかしAVFグループにおいては多くの学習者が1秒あたりの変化量を増加させたことから、視覚フィードバックによって、よりピッチの変化をつけて発音しようと意識づけられた可能性がある。

発音を指導するためのシステムは近年多く開発されているが、学習者にとってはもちろんのこと、指導する立場の人間にとっても使いやすく、学習管理がしやすいシステムが求められる。音声指導の重要性が高まる中で、発音練習における学習者及び指導者の負担を少しでも軽減させ、効率的に発音を向上させることができるよう、今後もシステムをさらに改善させ、実践的な研究を進めていきたい。

## 参考文献

- ベネッセ教育総合研究所(2015).「中高の英語指導に関する実態調査2015」, <http://berd.benesse.jp/global/research/detail.php?id=4776> (閲覧日: 2017年3月10日).
- ベネッセ教育総合研究所 (2015).「小学生の英語学習に関する調査」, <http://berd.benesse.jp/global/research/detail.php?id=4760> (閲覧日: 2017年3月10日).
- Boersma, P., & Weenink, D.(2009). Praat: Doing phonetics by computer. Retrieved February 11, 2016 from <http://www.fon.hum.uva.nl/Praat/>
- Dougiamas, M., & Taylor, P.(2003). Moodle: Using learning communities to create an open source course management system, Retrieved March 20, 2016 from <http://moodle.org/>
- Kashiwagi, A., & Snyder, M.(2010). Speech characteristics of Japanese speakers affecting American and Japanese listener evaluations. Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, 10, 1-14.
- 中川聖一 (2002).『日本人学生による読み上げ英語音声データベース (UME-ERJ)』東京: 大学共同利用機関法人情報・システム研究機構.
- Wilson, I. (2008). Using Praat and Moodle for teaching segmental and suprasegmental pronunciation. Proceedings of the 3rd international WorldCALL Conference: Using Technologies for Language Learning, 112-115.
- Yabuuchi, S., & Sato, H. (2001). Prosodic Characteristics of Japanese EFL Learners' Oral Reading Comparison between good and poor readers. Language Education & Technology, 38, 99-112.